Atmos. Chem. Phys. Discuss., 15, 26273–26313, 2015
www.atmos-chem-phys-discuss.net/15/26273/2015/
doi:10.5194/acpd-15-26273-2015

Atmospheric Chemistry and Physics Discussions

Ability of the 4-D-Var analysis of the GOSAT BESD XCO_2 retrievals to characterize atmospheric CO_2 at large and synoptic scales

S. Massart[1], A. Agustí-Panareda[1], J. Heymann[2], M. Buchwitz[2], F. Chevallier[3], M. Reuter[2], M. Hilker[2], J. P. Burrows[2], F. Hase[4], F. Desmet[5], D. G. Feist[6], and R. Kivi[7]

[1]European Centre for Medium-Range Weather Forecasts, Reading, UK
[2]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[3]Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, IPSL, Gif sur Yvette, France
[4]Karlsruhe Institute of Technology, IMK-ASF, Karlsruhe, Germany
[5]University of Antwerp, Antwerp, Belgium
[6]Max Planck Institute for Biogeochemistry, Jena, Germany
[7]Finnish Meteorological Institute, Arctic Research, Sodankylä, Finland

Received: 23 July 2015 – Accepted: 3 September 2015 – Published: 28 September 2015

Correspondence to: S. Massart (sebastien.massart@ecmwf.int)

Published by Copernicus Publications on behalf of the European Geosciences Union.

ACPD
15, 26273–26313, 2015

Abstract

This study presents results from the European Centre for Medium-Range Weather Forecasts (ECMWF) carbon dioxide (CO_2) analysis system where the atmospheric CO_2 is controlled through the assimilation of column-average dry-air mole fractions of CO_2 (XCO_2) from the Greenhouse gases Observing Satellite (GOSAT). The analysis is compared to a free run simulation and they are both evaluated against XCO_2 data from the Total Carbon Column Observing Network (TCCON). We show that the assimilation of the GOSAT XCO_2 product from the Bremen Optimal Estimation DOAS (BESD) algorithm during the year 2013 provides XCO_2 fields with an improved station-to-station bias deviation of 0.7 parts per million (ppm) compared to the free run (1.4 ppm) and an improved estimated precision of $\sim$ 1 ppm compared to the used GOSAT data (3.4 ppm). We also show that the analysis has skill for synoptic situations in the vicinity of frontal systems where the GOSAT retrievals are sparse due to cloud contamination. We finally computed the 10 day forecast from each analysis at 00:00 UTC. Compared to its own analysis the CO_2 forecast shows synoptic skill for the largest scale weather patterns even up to day 5 according to the anomaly correlation coefficient.

1 Introduction

Carbon in the atmosphere is mostly in the form of carbon dioxide (CO_2). Its amount is relatively small compared to the amount of carbon present in some of the other reservoirs like the ocean (Ciais et al., 2013). Being well mixed, the atmospheric CO_2 is nevertheless easier to monitor by the mean of in situ measurements than the carbon of some other reservoirs. The atmospheric reservoir varies as a result of changes in the surface fluxes to and from the atmosphere. Monitoring the atmospheric CO_2 therefore provides insight not only on the atmospheric CO_2 but potentially provides information about the surface fluxes. To improve the monitoring of atmospheric CO_2, one can combine atmospheric CO_2 measurements with a numerical model. This paper

ACPD
15, 26273–26313, 2015

describes such a system which has been developed for the Copernicus Atmosphere Monitoring Service (CAMS).

Rather than using the relatively sparse network of the surface air-sample measurements, we explore here the measurements from satellite sounders in order to have a more global picture of the atmospheric CO_2. To extract information on the CO_2 content in the atmosphere, passive atmospheric remote sounders measure in the thermal infrared (TIR) or the near infrared (NIR) in combination with the short wave infrared (SWIR).

The Atmospheric Infrared Sounder (AIRS) measuring in the TIR, detects thermal radiation emitted by the Earth's surface and the atmosphere (Chédin et al., 2003). The assimilation of the AIRS observed radiances was developed by Engelen et al. (2009) at the European Centre for Medium-Range Weather Forecasts (ECMWF) using a four-dimensional variational (4-D-Var). Their results showed the potential of data assimilation to constraint the atmospheric CO_2. They also showed the limitations of the assimilation of AIRS radiances, in particular due to the specific vertical sensitivity of the sounder. Because of the low thermal contrast between the Earth's surface and the air masses above, AIRS measurements have a limited sensitivity or no sensitivity to the lower troposphere and a higher sensitivity to the middle atmosphere. The signals of the CO_2 surface sources and sinks being the largest in the near-surface and lower troposphere, AIRS measurements were not able to capture these signals.

In contrast, column-average dry-air mole fractions of CO_2 (or XCO_2) with a high near-surface sensitivity are retrieved from NIR/SWIR measurements based on backscattered solar radiation. However, the NIR/SWIR measurements also have their limitations. They need sunlight and they are therefore limited to daytime observations. Sufficiently cloud-free conditions are also needed for accurate XCO_2 retrievals.

The aim of this study is to document the assimilation of XCO_2 products from NIR/SWIR measurements and how this impacts the simulated atmospheric CO_2 concentration. For that purpose, we assimilated the XCO_2 products derived from the NIR/SWIR spectra of the Greenhouse gases Observing Satellite (GOSAT,

ACPD
15, 26273–26313, 2015

www.gosat.nies.go.jp). The assimilation system is based on the ECMWF system of Engelen et al. (2009) which has lately evolved for CAMS in order to assimilate XCO_2 retrieved products instead of observed radiances (Massart et al., 2014).

The assimilation system provides an analysis of the atmospheric CO_2 concentration which is then integrated in time using a forecast model. The CO_2 forecast model used in this study is documented by Agustí-Panareda et al. (2014). In this model the CO_2 budget is not constrained and it may deviate from reality. The error in the CO_2 budget accumulates in the atmosphere and could lead to large global biases in the atmospheric CO_2. On the other hand, the strength of the CO_2 forecast model is its ability to provide a realistic CO_2 variability. The first objective of this study is to determine the quality of the CO_2 fields resulting from the assimilation of GOSAT XCO_2 data with a CO_2 forecast model where the CO_2 budget is not constrained.

The atmospheric CO_2 variability on a regional scale is related to synoptic events and frontal systems (Wang et al., 2007). These events are difficult to capture by the GOSAT measurements as the availability of the data is limited due to cloud contamination. The second objective of this study is to document if the assimilation helps improve the atmospheric CO_2 field for synoptic events despite the lack of measurements nearby.

Within CAMS, ECMWF is providing a CO_2 analysis based on the assimilation of the GOSAT XCO_2 data with a delay of 5 days behind real time. A ten day forecast is then issued from the analysis in order to provide the atmospheric CO_2 field in real time and for the next few days. The last objective of this study is to assess the quality of this forecast. The forecast quality, as a function of the lead time and the season, is evaluated against the analysis.

This paper is structured as follows. Section 2 introduces the data sets used in this study. Section 3 describes our atmospheric CO_2 simulations with and without assimilation of the GOSAT XCO_2 data, and how we compared them with independent measurements. Sections 4 to 6 present the global evaluation of our simulations, a case study and the evaluation of the CO_2 forecast based on the analysis. Finally, Sect. 7 presents our conclusions.

ACPD
15, 26273–26313, 2015

2 Data sets

In this study, we used two sets of data. The first one is the XCO_2 product obtained by the University of Bremen (UoB) and described in Sect. 2.1. The second one is the collection of measurements provided by the Total Carbon Column Observing Network (TCCON, Wunch et al., 2011a) and described in Sect. 2.2.

2.1 GOSAT XCO_2

The GOSAT satellite is a joint effort from the Japanese Aerospace Exploration Agency (JAXA), the National Institute for Environmental Studies (NIES) and the Japanese Ministry of the Environment (MOE) as part of the Global Change Observation Mission (GCOM) program of Japan. The GOSAT satellite was launched on 23 January 2009 and it carries the Thermal And Near-infrared Sensor for carbon Observations, which consists of a Fourier Transform Spectrometer (TANSO-FTS) and a Cloud and Aerosol Imager (TANSO-CAI).

In this study, we used XCO_2 retrieved from TANSO-FTS measurements of the upwelling radiance at the top of the atmosphere by the Bremen Optimal Estimation DOAS (BESD) algorithm of UoB. The BESD algorithm was initially developed to retrieve XCO_2 from nadir measurements of the SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) remote sensing spectrometer (Reuter et al., 2010, 2011). The BESD algorithm has been modified to also retrieve XCO_2 from GOSAT measurements. A detailed description of the GOSAT BESD algorithm can be found in Heymann et al. (2015). In brief the algorithm uses three fitting windows, the O_2-A band (12 920–13 195 cm^{-1}), a weak CO_2 absorption band (6170–6278 cm^{-1}) and a strong CO_2 band (4804–4896 cm^{-1}). An optimal estimation based inversion technique is used to derive the most probable atmospheric state from every individual GOSAT measurement using a priori knowledge. The BESD algorithm explicitly accounts for atmospheric scattering by clouds and aerosols which reduces potential

ACPD
15, 26273–26313, 2015

systematic biases. The scattering information is mainly obtained from the strong O_2-A and strong CO_2 absorption bands.

We used an inhomogeneous GOSAT BESD XCO_2 dataset in this study as the GOSAT BESD algorithm was still under development. This intermediate version of the GOSAT BESD XCO_2 data is referred to as MACC GOSAT BESD XCO_2 (MACC standing for Monitoring Atmospheric Composition and Climate, the precursor of CAMS). Nevertheless, from beginning of 2014 onwards, we have been assimilating in near real time the current version of the GOSAT BESD data (v01.00.02, Heymann et al., 2015).

The TANSO-FTS detector has a circular field of view of 10.5 km when projected on the Earth's surface (at exact nadir). It measured in 2013 in a mode with 3 measurements across track, the footprints being separated by ~ 263 km across track and ~ 283 km along track. The GOSAT satellite could operate in target mode resulting in a finer sampling distance. For these specific situations, we further thinned the observations on a $1^\circ \times 1^\circ$ grid by removing all the observations but one. This procedure avoids having several measurements in the same model grid cell during the assimilation. This thinning plus the characteristics of the instrument (measurement only during sunlit) and the processing of the level-2 data procedure (retrievals for clear-sky conditions and only over land) reduces the number of GOSAT XCO_2 data to about 100 per day. The assimilation window being 12 hours, this means that about 50 GOSAT XCO_2 data are assimilated each time.

The geographic distribution of these data is very much dependent on the season and the atmospheric conditions as illustrated by Fig. 1. For example, in July 2013 GOSAT BESD data are available up to 75° N, in October 2013 they are available up to 60° N. The reason for this is the quality filtering. Measurements under high solar zenith angle (SZA) conditions are more challenging to evaluate as the impact of atmospheric scattering becomes larger compared to low SZA conditions. Other data gaps are due to the strict cloud and quality filtering.

ACPD
15, 26273–26313, 2015

For the assimilation, the observation error covariances have to be specified. In this study, we assume that the observation errors are not correlated and the uncertainty of the individual (single ground pixel) XCO_2 retrieval is about 2 parts per million (ppm).

2.2 TCCON XCO_2

In 2014, the version GGG2014 of the TCCON data was released (http://tccon.ornl.gov/). Within the time period we are interested in (year 2013), data from 17 TCCON stations are available.

Not all the stations were used in this study. First we removed JPL 2011 and Pasadena/Caltech, as they are not background stations and are associated with significant representativeness error. We also removed Edwards/Dryden/Armstrong. This station started to report data from the middle of the year 2013, and we assumed that this was not long enough to provide information on the seasonal variation of the error in our simulations. Additionally, we removed Eureka from the list of stations as the site was reporting during only three days in 2013. This selection of the TCCON station made us retain 14 stations (Table 1).

Orléans had a specific treatment compared to the other stations. The averaging kernel was not specified in the GGG2014 release. So we decided to use the same information as for Lamont as advised in the previous release of the TCCON data (version GGG2012).

3 Experimental setup

We ran two model simulations for the year 2013. The first one is similar to the operational CAMS CO_2 forecast (Agustí-Panareda et al., 2013) and is referred to as the free run. This simulation is used as the reference to assess the impact of the assimilation of the GOSAT BESD XCO_2 data. The second simulation is the analysis in which the GOSAT XCO_2 data are assimilated and is referred to as the analysis. The config-

ACPD
15, 26273–26313, 2015

uration of both simulations is described in Sect. 3.1. The simulations were evaluated one against the other and also against the TCCON data. Section 3.2 introduces the methodology used to compared a simulation with the TCCON data.

3.1 Model simulations

The global simulations of the atmospheric CO_2 are performed within the Numerical Weather Prediction (NWP) framework of the Integrated Forecasting System (IFS). The CO_2 mass mixing ratio is directly transported within IFS as a tracer and is affected by surface fluxes. The transport is computed online and is updated each 12 h benefiting from the assimilation of all the operational observations within the IFS 4-D-Var assimilation system. The terrestrial biogenic carbon fluxes are also computed online by the carbon module of the land surface model (CTESSEL, Boussetta et al., 2013) while other prescribed fluxes are read from inventories (see Agustí-Panareda et al., 2014 for more details).

The ability of assimilating retrieval products from GOSAT was included in the IFS and is detailed in Massart et al. (2014) for the assimilation of methane data. The system used in this study is similar to the one of Massart et al. (2014) and is based on fixed background errors derived from the National Meteorological Center (NMC) method (Parrish and Derber, 1992). The standard deviation of the background error are constant on each model level and the constant value slowly increases from the upper troposphere to the lower troposphere with values from about 1 to about 5 ppm and rapidly increases to reach the value of about 40 ppm at the surface. The correlation of the background errors varies over the whole domain and in the vertical with a representative length scale of about 250 km. The system does not account for the spatial or temporal correlation between the errors of the observations.

We chose in this study to have a horizontal resolution of T255 on a Gaussian grid ($\sim 80\,\text{km} \times 80\,\text{km}$), and 60 vertical levels from the surface up to 0.1 hPa. This resolution is sufficient for resolving the large and synoptic scale horizontal structures of the atmospheric CO_2 fields.

ACPD
15, 26273–26313, 2015

3.2 Comparison with TCCON

To evaluate the quality of the model simulations (free run and analysis), we have extensively used the TCCON data in this study. The comparison is performed in the TCCON space using the TCCON a priori and averaging kernel information (see Appendix A for more details). In order to have a decomposition of the errors of the model column-average CO_2 against the TCCON measurement, we computed the quantities

$$\begin{aligned} \text{bias} &= \text{mean}\left(\hat{c} - \hat{c}^o\right), \\ \text{scatter} &= \text{std}\left(\hat{c} - \hat{c}^o\right), \end{aligned} \tag{1}$$

where mean is the simple average and std is the standard deviation over a sample of TCCON measurements grouped into $\hat{c}^o$ and of model equivalent values grouped into $\hat{c}$. The bias of Eq. (1) is consequently the sum of the bias of the model (in terms of column-average) and the bias of the TCCON retrieved column-average. Assuming that the TCCON bias is low compared to the model bais, the bias of Eq. (1) is a measure of the model bias. At the first order, the scatter of Eq. (1) is the sum of the random error of the model and the random error of the TCCON retrieved column-average (usually of the order of 0.5 ppm).

Additionally, we computed the correlation coefficient between $\hat{c}$ and $\hat{c}^o$. The statistics are computed for each station individually, but we also computed the mean of the individual station bias, known as the model offset following Heymann et al. (2015). Additionally, we computed the station-to-station bias deviation (or bias deviation) as the standard deviation of the station biases and we estimated the model precision as the mean of the station scatter.

The statistics of the simulations against the TCCON data have some gaps in time due to availability of the TCCON data. They are also valid only where the TCCON sites are located, i.e. 14 points distributed on the globe. To have a more global overview of the model bias and scatter against the TCCON data, we smoothed these statistics in time and space (see Appendix B for more details). In summary for the bias, we averaged

ACPD
15, 26273–26313, 2015

over a week period all the model–measurement differences for each TCCON site. We then fit the time evolution of the weekly bias with a function that combines a linear and a harmonic component for each station. The second step is an extrapolation in space. For each week, the weekly biases of every station are extrapolated using a quadratic function of latitude. This results in a Hovmöller diagram of the bias as a function of time and latitude. As similar process is applied for the scatter.

4 Global evaluation of the analysis

In this section we describe our evaluation of the quality of the XCO_2 delivered by the assimilation of the MACC GOSAT BESD data. As described previously, part of the evaluation is based on the comparison with the free run simulation. Thus, we first present the characteristics of the XCO_2 derived from the free run simulation when compared to the TCCON data. The second part presents the impact of the assimilation of the MACC GOSAT BESD XCO_2 comparing the XCO_2 from the analysis against the XCO_2 from the free run. Then, we discuss if the analysis represents an improvement compared to the free run in terms of statistics against the TCCON data. Finally, we discuss the merits of the analysis compared to the MACC GOSAT BESD data using once more the TCCON data as a reference.

4.1 Free run simulation vs. TCCON

When compared with the TCCON data, the free run simulation has a mean bias of −0.42 ppm (Table 2). However, station-to-station bias variations are large and span a range from 1.8 ppm at Darwin to −2.9 ppm at Białystok. The station-to-station bias deviation of the forecast is therefore large with a value of 1.44 ppm.

The variations of the bias as well as the seasonal cycle of the bias are highlighted on the Hovmöller diagram displayed in Fig. 2a. First, it shows that the initial condition of the free run has a positive bias of more than 2 ppm over the tropical region when compared

ACPD
15, 26273–26313, 2015

to the TCCON data. This bias is reduced during the spring and reappears the next summer. It reaches its highest values in autumn with more than 2 ppm. These results are slightly different from those of Agustí-Panareda et al. (2014) where the model bias was found to be more constant in the tropical region when comparing the background CO_2 in the marine boundary layer with the NOAA GLOBALVIEW-CO2. Here, the evaluation of the bias in the tropics is driven by the comparison with XCO2 measurements from the TCCON station of Ascension Island. For this station, the values of the bias from July to September result from the interpolation process as no measurements were reported during this period (Fig. S1 of the Supplement).

Contrary to the tropics, the initial condition of the free run has a negative bias at northern mid-latitudes and reaches almost 5 ppm at the latitude of Sodankylä (67° N) when compared to the TCCON XCO_2. This value is the result of the smoothing process as we do not have data for that period (Fig. 4a). The negative bias at these mid-latitudes is confirmed by the comparison with other stations like Park Falls where we have some data for the beginning of the year (Fig. 4b). The negative bias at mid-latitudes remains high during the whole year, with an absolute value generally greater than 1 ppm and larger at the end of spring, and in June and December. This can be explained by the fact that the model does not release enough CO_2 before and after the growing season, i.e. March to May and October to December, and by the fact that the onset of the CO_2 sink associated with the growing season starts too early in the model (Agustí-Panareda et al., 2014).

The precision of the forecast measured with the average scatter between the free run and the TCCON data is 1.4 ppm (Table 2). As for the bias, the scatter presents some variation in time and space (Fig. 3a). The Hovmöller diagram highlights the northern mid-latitudes during May–June–July when the scatter has its highest values of more than 1 ppm. This increase in the scatter is driven by the behaviour of the free run at Sodankylä whereas the simulation has a lot of variability when the measurements show no variability (Fig. 4a). Elsewhere, there is also an increase of the scatter between May and July which is the growing season. This increase could be explained by the issues

ACPD
15, 26273–26313, 2015

of the model during this season or by the higher variability of the atmospheric CO_2 during this season.

4.2 Analysis vs. free run

To assess the impact of the assimilation of the MACC GOSAT BESD XCO_2, we compared the evolution of the CO_2 total column from the analysis with the CO_2 total column from the free simulation. Figure 5 presents the Hovmöller diagram (time vs. latitude) of the difference. It shows that the first region where the analysis impacts the total column is the tropics. There, compared to the free run, the analysis continuously decreases XCO_2, up to 1 ppm in June and more that 2 ppm from September to December. The assimilation of the GOSAT data has consequently a positive impact as the free run has a positive bias in this region in autumn compared to the TCCON data.

The analysis also decreases XCO_2 over the southern extra tropics when compared to the free run. The decrease extends to the southern high latitudes even if we had no GOSAT data to assimilate in this region. This decrease results mainly from the transport of CO_2 from the equatorial region and mid-latitudes towards high latitudes. Unfortunately, there is no independent data available at southern high latitudes to assess the merits of the analysis there.

Even if some GOSAT data are assimilated in the northern mid-latitudes during the first months of the simulation, the analysis starts to differ significantly from the free run only from March onwards. In this region, north of 30° N, the analysis has higher values of XCO_2 than for the free run, with a difference of more than 2 ppm during the northern summer. Again, the assimilation of the GOSAT data is evident as the free run shows a strong negative bias there. Similar to the behaviour discussed for the southern high latitudes, the change in the CO_2 concentration at northern mid-latitudes is transported northward at higher latitudes. There is, nevertheless, a difference between the two hemispheres. For the Northern Hemisphere we have more data at high latitudes, especially during the summer when the northernmost GOSAT measurement goes up to 80° N.

ACPD
15, 26273–26313, 2015

4.3 Analysis vs. TCCON data

When compared with the TCCON data, the GOSAT BESD XCO_2 analysis has a mean bias of −0.47 ppm (Table 2). This is similar to the one of the free run (−0.42 ppm). But the bias is more constant in time for the analysis compared to the free run. For example, the slope of the fitting curve for the bias of the free run is for the investigated time period 2.08 ppm yr^{-1} for Lauder (Table S1 of the Supplement). It improves to 0.47 ppm yr^{-1} for the analysis (Table S2 of the Supplement and Fig. 4c).

By increasing XCO_2 in the northern mid-latitudes as discussed before, the analysis considerably reduces the bias. A residual seasonal cycle bias is still present with values usually in the range of 0 to 3 ppm (Fig. 2b). This could be explained by the fact that we correct the atmospheric state of CO_2 and not the CO_2 fluxes. During the seasons when the CO_2 fluxes are the main driver of the atmospheric CO_2, the optimisation of the atmospheric state may not be enough.

The analysis has a more constant bias in time than the free run, and is also more accurate in space with a station-to-station bias deviation which is largely reduced compared to the free run with a value of 0.66 ppm against 1.44 ppm (Table 2). The assimilation of the MACC GOSAT BESD XCO_2 thus helps improve significantly the accuracy of the model. The assimilation also helps improve the precision with a mean scatter improved by 17 %, with a value of 1.14 ppm. The scatter of the analysis is reduced for all TCCON stations compared to the free run except for Garmisch where the scatter is similar between the free run and the analysis. The Hovmöller diagram of the scatter shows that the main reduction is in the northern high latitudes in May (Fig. 3). In particular, the analysis shows less spurious variability than the free run at Sodankylä (Fig. 4a).

4.4 Analysis vs. MACC GOSAT BESD data

The analysis proves to be much more accurate and more precise than the free run when compared to the TCCON data. The analysis also fills the gaps in time and space

ACPD
15, 26273–26313, 2015

of the MACC GOSAT BESD data. In this section, we evaluate the analysis against the MACC GOSAT BESD data using one more time the TCCON data as a reference.

The MACC GOSAT BESD data were compared to the TCCON data using a geolocation criteria of 5° in space and a time window of ±2 h. For each GOSAT/TCCON pair, we added a correction to account for the usage of different a priori CO_2 profiles in the two products, before computing the difference. Moreover, we kept the stations where more than 30 GOSAT/TCCON pairs were found in order to have more robust statistical results. This procedure reduces the number of TCCON stations used in the comparison to 10 (Table 3).

For each GOSAT/TCCON pair, we extracted the CO_2 profile from the analysis at the same location and time of the GOSAT measurement before computing the difference between the model and the TCCON data. In this way, we have a fair comparison between the analysis and the MACC GOSAT BESD data with respect to the TCCON data.

The resulting subset of the analysis/TCCON differences has a different offset than the full dataset but a similar bias deviation and mean scatter (Tables 2 and 3). The difference in the offset is mainly due to a difference in the sampling between the subset and the full dataset over the Northern Hemisphere. Due to few or no pairs in Spring for the subset, the sampling misses the negative bias of the analysis there. Missing the negative bias of the analysis results in increasing the offset.

With a value just under 0.7 ppm, the station-to-station bias deviation of the analysis is almost half of the deviation of the MACC GOSAT BESD data (1.3 ppm, Table 3). With the successive improvements in the BESD algorithm, the deviation of the retrieved product decreased and is curently ∼ 0.4 ppm in the latest version of the product (Heymann et al., 2015). With a value of 1 ppm, the precision is more than 3 times less in the analysis than for the MACC GOSAT BESD data (3.4 ppm, and ∼ 2 ppm for the latest version of the product). The mean correlation coefficient is also higher in the analysis than in the data with a value of 0.8 compared 0.5.

ACPD
15, 26273–26313, 2015

The lower precision and station-to-station bias deviation of the analysis compared to the MACC GOSAT BESD dataset shows that the analysis is capable of smoothing the scatter of the satellite data. Moreover, the analysis is able to perform more than gap filling of the satellite data in time and space as it provides a more accurate and precise representation of the global atmospheric XCO_2 field.

5 Case study of a cold front over Park Falls

The CO_2 concentration could be strongly affected by frontal systems. As an illustration, such a situation occurred end of May 2013, close to the TCCON station of Park Falls, Wisconsin, USA, when a cold front is coming from the North-West. On 30 May, the XCO_2 measured at Park Falls at 08:21 LT is 399.350 ppm. About one day later, on 31 May, the XCO_2 dropped from 398.62 ppm at 08:15 LT to 395.97 ppm at 12:53 LT (Fig. 6, top panel). This sudden decrease of 2.65 ppm in less than 5 h occurs after the arrival of a cold front associated with a decrease of the surface pressure and a decrease of the temperature at 500 hPa (Fig. 6, lower panel).

The free run is able to capture the sudden decrease in XCO_2, highlighting the ability of the model for such a situation (Fig. 6, upper panel). The flow during this period is mainly a descent of cold air from Canada towards the Midwestern and Eastern US. This cold air mass is poorer in CO_2 than the background (Fig. 7e and b). When it moves towards Park Falls, it results in decreasing XCO_2 as observed and simulated, but the decrease in the free run is too strong compared to the measurements.

We have investigated if the assimilation of the GOSAT data could help improve the simulated evolution of the CO_2 concentration for such situations even if the number of BESD GOSAT data is limited in the vicinity of a frontal system due to the strict cloud filtering. Frontal systems are associated with clouds formed when moist air between the cold and warm fronts is lifted.

On May, 30 we have a few GOSAT measurements over the North and North-East region of North America (Fig. 7a). These measurements have the effect of increas-

ACPD
15, 26273–26313, 2015

ing the XCO_2 in this region (Fig. 7b–d). The cold air mass is then richer in CO_2 in the analysis compared to the free run, and when it moves towards Park Falls, the decrease is weaker and closer to the observed decrease. The assimilation of the GOSAT data, by correcting upstream the large scale structure and by improving the large scale atmospheric CO_2 gradient helps improve the simulation.

The XCO_2 decrease continues the next day on 1 June, in both simulations as the cold front continued its descent. Unfortunately, likely due to the presence of clouds, no TCCON measurements are available during this period to corroborate the simulated XCO_2 decrease.

6 Forecast based on the analysis

Within CAMS, we are receiving the GOSAT BESD data for a given day with a delay of 5 days behind real time. We are running the analysis for this day as soon as the data are received. We are finally running a 10 day forecast based on the resulting analysis.

In this section, we aim at evaluating the forecast as a function of the lead time by comparing the forecast to the analysis valid for the same time. This comparison informs us how long the information provided by the analysis lasts in the forecast. Assuming perfect transport and perfect surface fluxes, the analysis and the forecast (valid for the same time) should be similar given that the analysis corrects accurately the atmospheric concentration of CO_2. In practice, the differences we will observe between the analysis and the forecast would either come from the transport, from the surface fluxes or from the analysis.

To compare a forecast with the analysis valid for the same time, we computed the anomaly correlation coefficient (ACC) for the total column of CO_2 (see Appendix C for more details). The ACC can be regarded as a skill score relative to the climatology, the higher the better the forecast. In the framework of NWP, an ACC reaching 50 % corresponds to forecasts for which the error is the same as for a forecast based on

ACPD
15, 26273–26313, 2015

a climatological average. An ACC of about 80 % indicates some skill in forecasting large-scale synoptic patterns.

We computed the ACC for each month individually as we know that the surface fluxes, drivers of the difference between the forecast and the analysis, have a strong seasonal cycle. We also computed it for different domains (globe, tropics and mid to high latitudes) and for several forecast lead times, from 12 h up to 10 days. We found that the ACC is globally more than 90 % for day 3 and almost always more than 85 % for day 5 for each single month (Fig. 8a). This means that the forecast for today based on the analysis of 5 days ago shows the same large-scale synoptic XCO_2 patterns as the analysis. The information of the analysis therefore lasts long enough in the forecast to provide a good quality forecast (compared to the analysis). The information lasts longer in the tropics than in the Northern Hemisphere and slightly longer in the Northern Hemisphere than in the Southern Hemisphere (Fig. 8b to d). This difference between the two hemispheres may reflect that variability is much weaker in the Southern Hemisphere.

For longer term forecasts, globally, there are two particular months for which the ACC decreases faster than the others, i.e. July and December. For example, for these two months the ACC at day 5 is similar to the ACC at day 10 for October. This means that for July and December, the medium range XCO_2 forecast should be used more carefully. For July, the drop in skill occurs mainly over the Northern Hemisphere. The main reason is that the CO_2 fluxes are an even more important driver of the CO_2 concentration for this month. To better understand the impact of the surface fluxes, let us assume that in July we have too little release of CO_2 in the atmosphere in the model over the Northern Hemisphere (as confirmed by Fig. 2a). In the meantime, the analysis increases the CO_2 concentration helped by the GOSAT BESD data (Fig. 5). But the next forecast will decrease again the CO_2 concentration due to the underestimation of the CO_2 fluxes. This opposition between the analysis and the forecast explains the reduction in skill.

ACPD
15, 26273–26313, 2015

The global drop in skill for December is not directly related to a particular region as for July. It is nonetheless the second worst month for the tropics (after January) and the third worst for the Northern Hemisphere (together with September). Over the tropics during the winter, the reduction in skill is due to the opposite effect as for July over the Northern Hemisphere: the CO_2 fluxes are important and there is too much release of CO_2 in the atmosphere in the model. For these situations when the CO_2 fluxes are the main driver of the atmospheric CO_2, the only solution to improve the skill would be to optimise the CO_2 fluxes together with the CO_2 initial conditions.

7 Conclusions

The Copernicus Atmosphere Monitoring Service (CAMS) greenhouse gases data assimilation within the Numerical Weather Prediction (NWP) framework of the Integrated Forecasting System (IFS) is designed to correct the atmospheric concentration of CO_2 instead of the surface fluxes in order to constrain the atmospheric CO_2. This requires the use of a short assimilation window so the error related to the model errors could be neglected. In the case of atmospheric CO_2, model errors are related to potentially inaccurate surface fluxes or transport.

This article demonstrates the benefit of the assimilation of the XCO_2 data derived from the Greenhouse gases Observing Satellite (GOSAT) by intermediate versions of the Bremen Optimal Estimation DOAS (BESD) algorithm of the University of Bremen (UoB). The assimilation of the GOSAT BESD XCO_2 provides a CO_2 analysis which was compared to a free run forecast where the CO_2 concentration is not constrained by any CO_2 observation. The comparison was one year long (year 2013) and both simulations (analysis and free run) were evaluated against the measurements from the Total Carbon Column Observing Network (TCCON). We showed that the free run has a negative bias at northern mid-latitudes and a large positive bias in the tropical region with strong seasonal variations in both regions. These results are similar to

ACPD
15, 26273–26313, 2015

those obtained with the same model and a similar configuration by Agustí-Panareda et al. (2014), where the causes of the bias are well detailed.

The analysis significantly reduces these biases without completely removing them with a remaining mean bias of −0.5 ppm compared to the TCCON data. However, the accuracy estimated with the station-to-station bias deviation is 0.7 ppm. This represents a large improvement compared to the free run for which the accuracy is 1.4 ppm. The precision of the analysis estimated with the mean scatter is 1.1 ppm, sightly better than for the free run with a value of 1.4 ppm.

The analysis was compared to the assimilated MACC GOSAT BESD data using again the TCCON data as a reference. This comparison showed that the analysis has a lower station-to-station bias deviation than the assimilated data (0.7 ppm to be compared to 1.3 ppm). The precision is much better for the analysis with a scatter of 1. ppm, while the assimilated data have a scatter of 3.4 ppm. The CO_2 analysis is consequently an alternative to the XCO_2 GOSAT data as it provides a lower bias and better precision XCO_2 product which, moreover, does not have spatial and temporal gaps.

The pre-operational CAMS CO_2 analysis is currently assimilating in near real time the most recent version of the GOSAT BESD data presented by Heymann et al. (2015). These data have an improved bias deviation (∼ 0.4 ppm) and an improved precision (∼ 2 ppm) compared to the ones used in this study. The near real time CAMS CO_2 analysis should therefore have an improved station-to-station bias deviation and precision than the analysis presented in this paper.

We corrected the atmospheric concentration by only constraining the atmospheric concentration and not the surface fluxes. When and where the surface flux is a significant driver of the atmospheric concentration and if the assimilated data are not good enough or not numerous enough (in time and space), then constraining only atmospheric CO_2 does not allow to counter-balance the error in the surface flux. The next step is to further improve the carbon module CTESSEL in order to reduce the bias of the model. Another long term solution would be to constrain the surface flux at the same time as the concentration.

ACPD
15, 26273–26313, 2015

One strength of the CO_2 model used in this study is its ability to represent CO_2 variations associated with small scale weather situations (Agustí-Panareda et al., 2014). By correcting the large scale XCO_2 patterns and removing part of the model bias, we showed with a case study that the analysis is able to better represent the CO_2 variations associated with these situations. The horizontal resolution of this study is about 80 km × 80 km while the horizontal resolution of the pre-operational analysis we run daily is 40 km × 40 km. One should expect an even better representation of the CO_2 variability in this analysis. In the future, the resolution could be increased even further toward the ECMWF operational resolution of about 16 km × 16 km.

Despite the remaining bias in the analysis, the quality of the analysis is sufficient enough to assess the quality of the forecast as a function of the lead time. We showed that the forecast for day 3 and day 5 which will be the valid forecast range for today's forecast has an anomaly correlation error of 90 and 85 % respectively. This means that we are providing a CO_2 forecast with accurate synoptic features for today and for the next days. With a good representation of the variability and a bias mostly under 1 ppm, the CAMS atmospheric CO_2 proves to become a useful product, for example, for planning a measurement campaign. It could also be used as the a priori in the satellite retrieval algorithms or be used to evaluate the retrieval products from the Orbiting Carbon Observatory-2 (OCO-2, oco.jpl.nasa.gov).

Appendix A: Comparing the model against TCCON

For the comparison with the TCCON data, one has to account for the a priori information used in the retrieval that links $\hat{c}^o$, the TCCON retrieved XCO_2 to $\boldsymbol{x}^t$, the true (unknown) CO_2 profile (Wunch et al., 2011b),

$$\hat{c}^o = c^b + \boldsymbol{a}^\mathsf{T}\left(\boldsymbol{x}^t - \boldsymbol{x}^b\right) + \varepsilon\,, \tag{A1}$$

ACPD
15, 26273–26313, 2015

where $\boldsymbol{x}^b$ is an a priori profile of CO_2, $\boldsymbol{a}$ is a vector resulting from the product of the averaging kernel matrix with a dry-pressure weighting function vector (for the vertical integration), c^b is the averaged-column computed from $\boldsymbol{x}^b$, and ε is the error in the retrieved column. This error includes the random and systematic errors in the measured signal and in the retrieval algorithm.

To compare the model with the TCCON retrieved value, we used the same a priori information, so that the model profile $\boldsymbol{x}$ is converted to a column $\hat{c}$ by

$$\hat{c} = c^b + \boldsymbol{a}^{\mathrm{T}} \left(\boldsymbol{x} - \boldsymbol{x}^b \right) . \tag{A2}$$

The comparison between the simulation and TCCON occurs in the observation space with the difference between the model column $\hat{c}$ of Eq. (A2) and the TCCON column $\hat{c}^o$ of Eq. (A1),

$$\hat{c} - \hat{c}^o = \boldsymbol{a}^{\mathrm{T}} \left(\boldsymbol{x} - \boldsymbol{x}^{\mathrm{t}} \right) - \varepsilon . \tag{A3}$$

Let us define $\eta = \boldsymbol{a}^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^{\mathrm{t}})$ as the model error in terms of the column. It accounts for numerous errors, for example, the errors directly linked to the model processes like the transport, the errors in the surface fluxes, the representativeness error and the error due to the assimilation of the GOSAT XCO_2 data for the analysis. The difference between the smooth model column $\hat{c}$ and the TCCON column $\hat{c}^o$ is therefore the sum of the model error η and the error in the retrieved column ε.

To compute the model column $\hat{c}$ of Eq. (A2) equivalent to each TCCON measurement, we extracted the two model profiles which are closest to the measurement time and which are in the nearest grid point of the station where the measurement was realised. Then we did a time interpolation of the two profiles in order to have the model profile at the same time as the measurement. Eventually, we computed the column according to Eq. (A2).

ACPD
15, 26273–26313, 2015

Appendix B: Smoothing the statistics against TCCON

In order to have a more global view of the bias and the scatter of a simulation against the data from the TCCON network, we have developed and used a two-step algorithm. The first step consists in computing the statistics (bias and the standard deviation) for each week of 2013 and for each TCCON station when the in situ data are available. The weekly statistics are then interpolated in time using a function described in the following Sect. B1. This allows one to fill in the gaps in time when no data are available. We therefore have a value for the bias at each station and for each week. For the second step, we compute a quadratic function of latitude that best fits the interpolated biases for each week (Sect. B2).

B1 Time smoothing

For each TCCON station we aggregate per week the measurements and their equivalent from the simulation. For each week when more than 10 measurements are available, we compute the average and standard deviation of the difference between every model value and observation, in the observation space. The averaged difference (or bias) is then interpolated in time t with the function $\widetilde{b}(t)$ that combines a linear growth and a harmonic component,

$$\widetilde{b}(t) = a\,t + b + \alpha \sin\left(\frac{t}{\tau_1} + \varphi\right) + \beta \sin\left(\frac{t}{\tau_2} + \varphi\right). \quad \text{(B1)}$$

a, b, α, β and φ are the parameters of the function $\widetilde{b}(t)$ obtained by an optimisation procedure that minimises the distance between $\widetilde{b}(t)$ and the raw bias. τ_1 is chosen to be 6 months and τ_2 3 months. The form of the function of Eq. (B1) thus gives a linear growing bias and allows seasonal variations. A similar function is used for the standard deviation.

ACPD
15, 26273–26313, 2015

B2 Spatial smoothing

The time smoothing allows us to fill in the gaps in the time series of the bias for each station, when for one week we do not have any measurement to compare with. Following Bergamaschi et al. (2009), we then compute for each week the best fit of the interpolated biases with a quadratic function of latitude $\hat{b}$,

$$\hat{b}(\phi) = a\phi^2 + b\phi + c\,, \tag{B2}$$

with ϕ is the sine of the latitude. a, b and c are obtained by an optimisation procedure that minimises the distance between $\hat{b}$ and the weekly interpolated biases. A similar function is used for the standard deviation.

B3 Discussion

With a root mean square error (RMSE) mostly under 0.7 ppm and a correlation mostly over 0.8, the smoothed bias matches well with the weekly bias (Table S1 of the Supplement). The Hovmöller diagram can be considered as an accurate representation of the overall bias.

Compared to the bias, the fit between the time series of the weekly scatter and the regression is not as good for the scatter. The correlation coefficient is mostly between 0.5 and 0.7 (Table S1 of the Supplement).

Appendix C: Anomaly correlation coefficient

The anomaly correlation coefficient ACC between the forecast f and the analysis a is computed using the climatology c by

$$\mathrm{ACC} = \frac{\overline{(f-c)(a-c)}}{\sqrt{\overline{(f-c)^2}\,\overline{(a-c)^2}}}\,, \tag{C1}$$

ACPD
15, 26273–26313, 2015

where the overline, is the spatial and temporal average. For example, for the forecast range 24 hr, we take the XCO_2 fields from all the 24 hr forecasts for a given month, all the analyses valid for the same time, and a fixed climatology for this month.

The climatology is based on a free run simulation using the optimised CO_2 surface fluxes from Chevallier et al. (2010) which simulated the years from 2003 to 2012. For each month, we compute the average over the 10 years of the simulation, rescaling the mean so that the mean is the same as for the analysis, avoiding by this procedure the issue of the increase in CO_2 over time. The two dimensional climatology field for XCO_2 for the month m is

$$c(m) = \frac{1}{10} \sum_{y=2003}^{2012} \frac{1}{n(y,m)} \sum_{d=1}^{n(y,m)} \left[\Sigma(y,m,d) - \overline{\Sigma}(y,m,d) \right] + \overline{\Sigma}_{\mathrm{an}}(m) \,, \tag{C2}$$

where y is the year, n the number of days for the year y and the month m, d is an index for the day, $\Sigma(y,m,d)$ is the XCO_2 field from the simulation for the year y, the month m and the day d, $\overline{\Sigma}$ is a spatial average of Σ and $\overline{\Sigma}_{\mathrm{an}}(m)$ is the spatial and temporal average of the XCO_2 fields from the analysis for the month m (and the year 2013).

The Supplement related to this article is available online at doi:10.5194/acpd-15-26273-2015-supplement.

Author contributions. S. Massart designed and carried out the experiments with the help of A. Agustí-Panareda and advice from F. Chevallier, J. Heymann and M. Buchwitz. J. Heymann, M. Reuter, M. Hilker, M. Buchwitz and J. P. Burrows have been responsible for the design and operation of the BESD GOSAT XCO_2 retrieval algorithm. S. Massart prepared the manuscript with contributions from A. Agustí-Panareda, J. Heymann, M. Buchwitz, F. Chevallier, M. Reuter, M. Hilker, J.P. Burrows, and F. Hase. F. Desmet is the co-investigator of the La Réunion TCCON station. D. G. Feist is the PI of the Ascension TCCON station. F. Hase is the PI of the Karlsruhe TCCON station. R. Kivi is the PI of the Sodankylä TCCON station.

Acknowledgements. This study was funded by the European Commission under the European Union's Horizon 2020 programme. The development of the GOSAT BESD algorithm received

ACPD
15, 26273–26313, 2015

funding from the European Space Agency (ESA) Greenhouse Gases Climate Change Initiative (GHG-CCI). TCCON data were obtained from the TCCON Data Archive, hosted by the Carbon Dioxide Information Analysis Center (CDIAC) – http://tccon.onrl.gov. The authors are grateful to Marijana Crepulja for the acquisition of the BESD GOSAT data at ECMWF and the preparation of the data for the assimilation. The authors would like to acknowledge Paul Wennberg, PI of the Lamont and Park Falls TCCON stations.

References

Agustí-Panareda, A., Massart, S., Boussetta, S., Balsamo, G., Beljaars, A., Chevallier, F., Engelen, R., Peuch, V.-H., and Razinger, M.: The new MACC-II CO_2 forecast, ECMWF Newsletter, 135, 8–13, 2013. 26280

Agustí-Panareda, A., Massart, S., Chevallier, F., Boussetta, S., Balsamo, G., Beljaars, A., Ciais, P., Deutscher, N. M., Engelen, R., Jones, L., Kivi, R., Paris, J.-D., Peuch, V.-H., Sherlock, V., Vermeulen, A. T., Wennberg, P. O., and Wunch, D.: Forecasting global atmospheric CO_2, Atmos. Chem. Phys., 14, 11959–11983, doi:10.5194/acp-14-11959-2014, 2014. 26277, 26281, 26284, 26292, 26293

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A., and Levin, I.: Inverse modeling of global and regional CH_4 emissions using SCIAMACHY satellite retrievals, J. Geophys. Res.-Atmos., 114, D22301, doi:10.1029/2009JD012287, 2009. 26296

Blumenstock, T., Hase, F., Schneider, M., García, O. E., and Sepúlveda, E.: TCCON data from Izana, Tenerife, Spain, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.izana01.R0/1149295, 2014. 26303

Boussetta, S., Balsamo, G., Beljaars, A., Agusti-Panareda, A., Calvet, J.-C., Jacobs, C., van den Hurk, B., Viterbo, P., Lafont, S., Dutra, E., Jarlan, L., Balzarolo, M., Papale, D., and van der Werf, G.: Natural carbon dioxide exchanges in the ECMWF integrated forecasting system: implementation and offline validation, J. Geophys. Res.-Atmos., 118, 1–24, doi:10.1002/jgrd.50488, 2013. 26281

Chédin, A., Saunders, R., Hollingsworth, A., Scott, N., Matricardi, M., Etcheto, J., Clerbaux, C., Armante, R., and Crevoisier, C.: The feasibility of monitoring CO_2 from high-resolution in-

ACPD
15, 26273–26313, 2015

frared sounders, J. Geophys. Res.-Atmos., 108, 4064, doi:10.1029/2001JD001443, 2003. 26276

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, J. Geophys. Res.-Atmos., 115, D21307, doi:10.1029/2010JD013887, 2010. 26297

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Quéré, C. L., Myneni, R., Piao, S., and Thornton, P.: Climate change 2013: the physical science basis, in: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., chap. Carbon and Other Biogeochemical Cycles, Cambridge University Press, Cambridge, UK and New York, NY, USA, 465–570, 2013. 26275

De Maziere, M., Desmet, F., Hermans, C., Scolas, F., Kumps, N., Metzger, J.-M., Duflot, V., and Cammas, J.-P.: TCCON data from Reunion Island (La Reunion), France, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.reunion01.R0/1149288, 2014. 26303

Deutscher, N., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., Grupe, P., and Katrynski, K.: TCCON data from Bialystok, Poland, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.bialystok01.R0/1149277, 2014. 26303

Engelen, R. J., Serrar, S., and Chevallier, F.: Four-dimensional data assimilation of atmospheric CO_2 using AIRS observations, J. Geophys. Res.-Atmos., 114, D03303, doi:10.1029/2008JD010739, 2009. 26276, 26277

Feist, D. G., Arnold, S. G., John, N., and Geibel, M. C.: TCCON data from Ascension Island, Saint Helena, Ascension and Tristan da Cunha, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory,

ACPD
15, 26273–26313, 2015

Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.ascension01.R0/1149285, 2014. 26303

Griffith, D. W. T., Deutscher, N., Velazco, V. A., Wennberg, P. O., Yavin, Y., Aleks, G. K., Washenfelder, R., Toon, G. C., Blavier, J.-F., Murphy, C., Jones, N., Kettlewell, G., Connor, B., Macatangay, R., Roehl, C., Ryczek, M., Glowacki, J., Culgan, T., and Bryant, G.: TCCON data from Darwin, Australia, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.darwin01.R0/1149290, 2014a. 26303

Griffith, D. W. T., Velazco, V. A., Deutscher, N., Murphy, C., Jones, N., Wilson, S., Macatangay, R., Kettlewell, G., Buchholz, R. R., and Riggenbach, M.: TCCON data from Wollongong, Australia, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.wollongong01.R0/1149291, 2014b. 26303

Hase, F., Blumenstock, T., Dohe, S., Groß, J., and Kiel, M.: TCCON data from Karlsruhe, Germany, Release GGG2014R1, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.karlsruhe01.R1/1182416, 2014. 26303

Heymann, J., Reuter, M., Hilker, M., Buchwitz, M., Schneising, O., Bovensmann, H., Burrows, J. P., Kuze, A., Suto, H., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kawakami, S., Kivi, R., Morino, I., Petri, C., Roehl, C., Schneider, M., Sherlock, V., Sussmann, R., Velazco, V. A., Warneke, T., and Wunch, D.: Consistent satellite XCO_2 retrievals from SCIAMACHY and GOSAT using the BESD algorithm, Atmos. Meas. Tech. Discuss., 8, 1787–1832, doi:10.5194/amtd-8-1787-2015, 2015. 26278, 26279, 26282, 26287, 26292

Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Taura, C., Fukamachi, T., and Sakashita, M.: TCCON data from Saga, Japan, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.saga01.R0/1149283, 2014. 26303

Kivi, R., Heikkinen, P., and Kyro, E.: TCCON data from Sodankyla, Finland, Release GGG2014R0., TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.sodankyla01.R0/1149280, 2014. 26303

Massart, S., Agusti-Panareda, A., Aben, I., Butz, A., Chevallier, F., Crevoisier, C., Engelen, R., Frankenberg, C., and Hasekamp, O.: Assimilation of atmospheric methane products into the

ACPD
15, 26273–26313, 2015

MACC-II system: from SCIAMACHY to TANSO and IASI, Atmos. Chem. Phys., 14, 6139–6158, doi:10.5194/acp-14-6139-2014, 2014. 26277, 26281

Parrish, D. and Derber, J.: National Meteorological Center's spectral statistical interpolation analysis system, Mon. Weather Rev., 120, 1747–1763, 1992. 26281

Reuter, M., Buchwitz, M., Schneising, O., Heymann, J., Bovensmann, H., and Burrows, J. P.: A method for improved SCIAMACHY CO_2 retrieval in the presence of optically thin clouds, Atmos. Meas. Tech., 3, 209–232, doi:10.5194/amt-3-209-2010, 2010. 26278

Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J. P., Connor, B. J., Deutscher, N. M., Griffith, D. W. T., Heymann, J., Keppel-Aleks, G., Messerschmidt, J., Notholt, J., Petri, C., Robinson, J., Schneising, O., Sherlock, V., Velazco, V., Warneke, T., Wennberg, P. O., and Wunch, D.: Retrieval of atmospheric CO_2 with enhanced accuracy and precision from SCIAMACHY: validation with FTS measurements and comparison with model results, J. Geophys. Res.-Atmos., 116, D04301, doi:10.1029/2010JD015047, 2011. 26278

Sherlock, V. B., Connor, Robinson, J., Shiona, H., Smale, D., and Pollard, D.: TCCON data from Lauder, New Zealand, 125HR, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.lauder02.R0/1149298, 2014. 26303

Sussmann, R. and Rettinger, M.: TCCON data from Garmisch, Germany, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.garmisch01.R0/1149299, 2014. 26303

Wang, J.-W., Denning, A. S., Lu, L., Baker, I. T., Corbin, K. D., and Davis, K. J.: Observations and simulations of synoptic, regional, and local variations in atmospheric CO_2, J. Geophys. Res.-Atmos., 112, D04108, doi:10.1029/2006JD007410, 2007. 26277

Warneke, T., Messerschmidt, J., Notholt, J., Weinzierl, C., Deutscher, N., Petri, C., Grupe, P., Vuillemin, C., Truong, F., Schmidt, M., Ramonet, M., and Parmentier, E.: TCCON data from Orleans, France, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.orleans01.R0/1149276, 2014. 26303

Wennberg, P. O., Roehl, C., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., Allen, N., and Ayers, J.: TCCON data from Park Falls, Wisconsin, USA, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Infor-

ACPD
15, 26273–26313, 2015

mation Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.parkfalls01.R0/1149161, 2014a. 26303

Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., Allen, N., Dowell, P., Teske, K., Martin, C., and Martin, J.: TCCON data from Lamont, Oklahoma, USA, Release GGG2014R0, TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, doi:10.14291/tccon.ggg2014.lamont01.R0/1149159, 2014b. 26303

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The total carbon column observing network, Philos. T. Roy. Soc. A, 369, 2087–2112, doi:10.1098/rsta.2010.0240, 2011a. 26278

Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of CO_2 total columns from space, Atmos. Chem. Phys., 11, 12317–12337, doi:10.5194/acp-11-12317-2011, 2011b. 26293

ACPD
15, 26273–26313, 2015

Table 1. List on the used TCCON stations ordered by latitude from North to South.

Site	Lat	Lon	Starting date	Reference
Sodankylä (sodankyla01)	67.37	26.63	6 Feb 2009	Kivi et al. (2014)
Białystok (bialystok01)	53.23	23.02	1 Mar2009	Deutscher et al. (2014)
Karlsruhe (karlsruhe01)	49.10	8.44	19 Apr 2010	Hase et al. (2014)
Orléans (orleans01)	47.97	2.11	29 Aug 2009	Warneke et al. (2014)
Garmisch (garmisch01)	47.48	11.06	16 Jul 2007	Sussmann and Rettinger (2014)
Park Falls (parkfalls01)	45.94	−90.27	26 May 2004	Wennberg et al. (2014a)
Lamont (lamont01)	36.60	−97.49	6 Jul 2008	Wennberg et al. (2014b)
Saga (saga01)	33.24	130.29	28 Jul 2011	Kawakami et al. (2014)
Izaña (izana01)	28.30	−16.48	18 May 2007	Blumenstock et al. (2014)
Ascension Island (ascension01)	−7.92	−14.33	22 May 2012	Feist et al. (2014)
Darwin (darwin01)	−12.43	130.89	28 Aug 2005	Griffith et al. (2014a)
Reunion Island (reunion01)	−20.90	55.49	6 Oct 2011	De Maziere et al. (2014)
Wollongong (wollongong01)	−34.41	150.88	26 Jun 2008	Griffith et al. (2014b)
Lauder 125HR (lauder02)	−45.05	169.68	2 Feb 2010	Sherlock et al. (2014)

ACPD
15, 26273–26313, 2015

Table 2. Statistics of the XCO_2 difference between the simulations (free run and analysis) and the average hourly TCCON data (model-TCCON): mean difference (bias, in ppm), standard deviation (scatter, in ppm) and correlation coefficient (r). Also shown are the mean and deviation of the stations bias, the mean scatter and the mean r (last two rows). The second column (N) is the number of data used for computing the statistics.

		Free run			Analysis		
Site	N	Bias	Scatter	r	Bias	Scatter	r
Sodankylä	1324	−1.71	1.41	0.91	−0.63	1.39	0.91
Białystok	933	−2.91	2.04	0.81	−1.78	1.86	0.79
Karlsruhe	595	−2.34	1.77	0.80	−1.42	1.57	0.82
Orléans	577	−0.58	1.37	0.83	−0.04	1.20	0.89
Garmisch	753	−0.86	1.59	0.81	−0.26	1.59	0.80
Park Falls	1603	−1.60	2.04	0.82	−0.57	1.39	0.91
Lamont	1973	−0.17	2.06	0.63	−0.00	1.27	0.83
Saga	511	−1.26	1.46	0.80	−0.75	1.18	0.86
Izaña	276	0.28	0.76	0.90	0.41	0.58	0.95
Ascension Island	592	2.32	1.01	0.35	0.72	0.98	0.31
Darwin	2175	1.80	1.15	0.81	0.18	1.04	0.81
Reunion Island	1105	0.56	0.70	0.83	−0.75	0.55	0.84
Wollongong	1456	0.60	0.98	0.79	−0.75	0.88	0.77
Lauder	1005	0.06	0.78	0.89	−0.96	0.53	0.88
Mean	14	−0.42	1.37	0.78	−0.47	1.14	0.81
Deviation	14	1.44	–	–	0.66	–	–

ACPD
15, 26273–26313, 2015

Table 3. Statistics of the XCO_2 differences between the MACC GOSAT BESD dataset and the average hourly TCCON data (left block, GOSAT-TCCON) or the analysis and the average hourly TCCON data (right block, model-TCCON): mean difference (bias, in ppm), standard deviation (scatter, in ppm) and correlation coefficient (r). The analysis has been sampled similarly to the GOSAT dataset in time and space. Also shown are the mean and deviation of the stations bias, the mean scatter and the mean r (last two rows). The second column (N) is the number of data points used for computing the statistics.

		MACC GOSAT dataset			Analysis		
Site	N	Bias	Scatter	r	Bias	Scatter	r
Sodankylä	90	−0.26	4.50	0.39	0.24	1.41	0.92
Białystok	73	0.23	3.35	0.25	1.00	1.79	0.35
Karlsruhe	94	0.62	2.74	0.53	0.16	0.79	0.86
Orléans	52	0.20	2.44	0.34	1.29	0.57	0.84
Garmisch	76	1.64	3.10	0.55	1.17	1.06	0.77
Park Falls	63	1.50	3.22	0.71	−0.08	1.03	0.95
Lamont	340	−1.01	4.05	0.57	0.05	1.01	0.91
Saga	61	0.40	2.95	0.76	0.14	0.88	0.90
Darwin	234	−1.27	3.37	0.42	−0.11	0.81	0.84
Wollongong	221	−3.03	3.86	0.31	−1.17	0.95	0.79
Mean	10	−0.10	3.36	0.48	0.27	1.03	0.81
Deviation	10	1.32	–	–	0.69	–	–

ACPD
15, 26273–26313, 2015

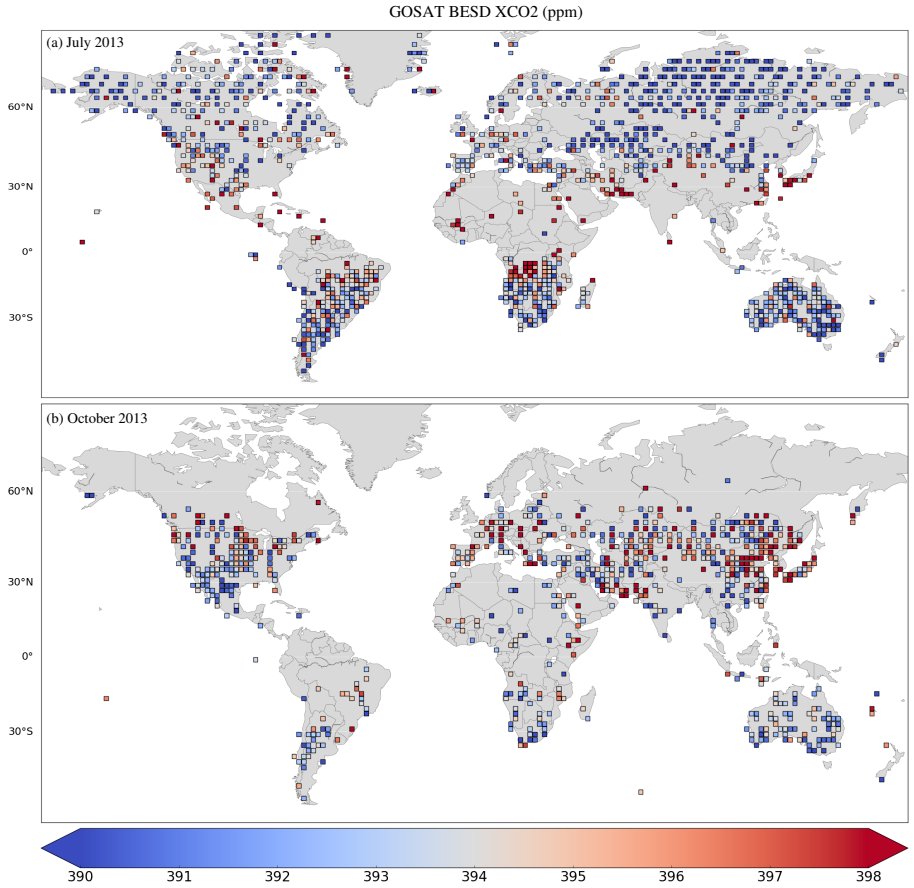

Figure 1. Example of the distribution of the assimilated GOSAT XCO_2 data: July 2013 (top panel, about 3400 data) and October 2013 (bottom, about 1270 data). The monthly data are here aggregated on a $2^\circ \times 2^\circ$ grid and averaged. The blue/red represents the low/high averaged XCO_2 values in ppm.

ACPD
15, 26273–26313, 2015

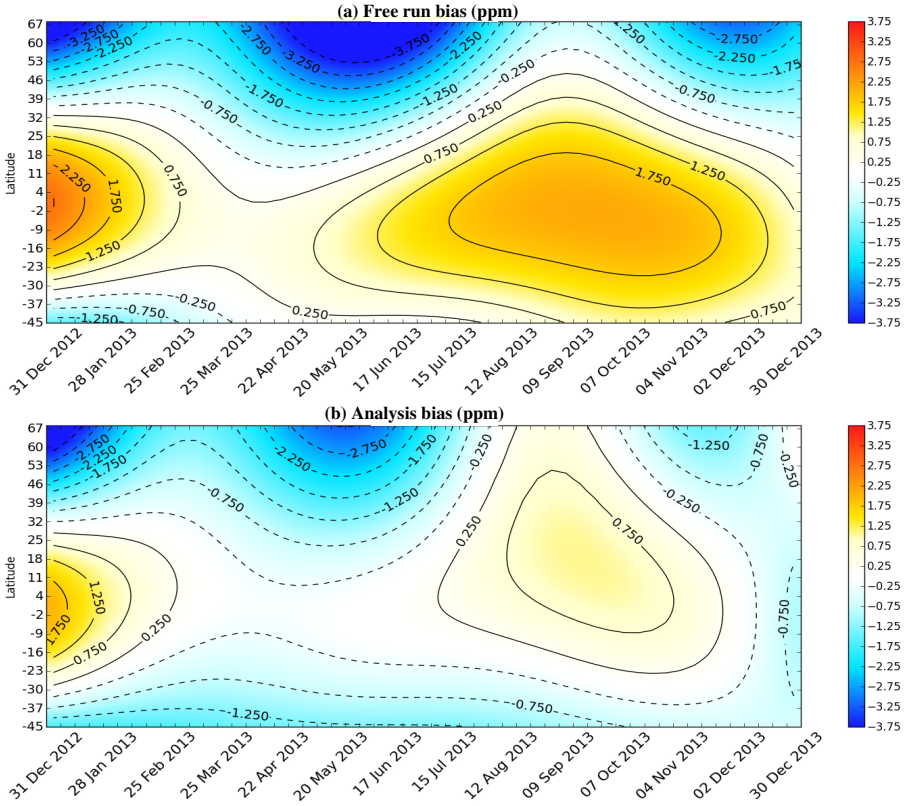

Figure 2. Hovmöller diagram (latitude vs. time) of the smoothed bias (in ppm, negative/positive in blue/red) of the simulated XCO_2 against the data of the TCCON network. Top: free run simulation. Bottom: analysis.

ACPD
15, 26273–26313, 2015

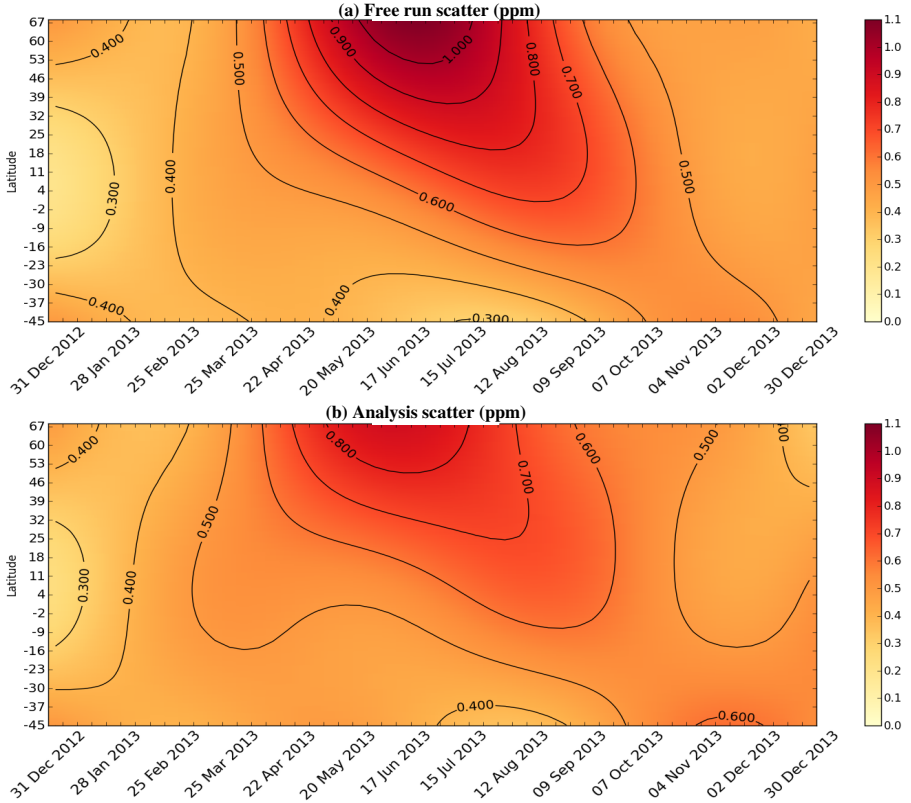

Figure 3. Same as Fig. 2 but for the standard deviation and yellow/red for low/high values.

ACPD
15, 26273–26313, 2015

4-D-Var analysis of the GOSAT BESD XCO_2 retrievals

S. Massart et al.

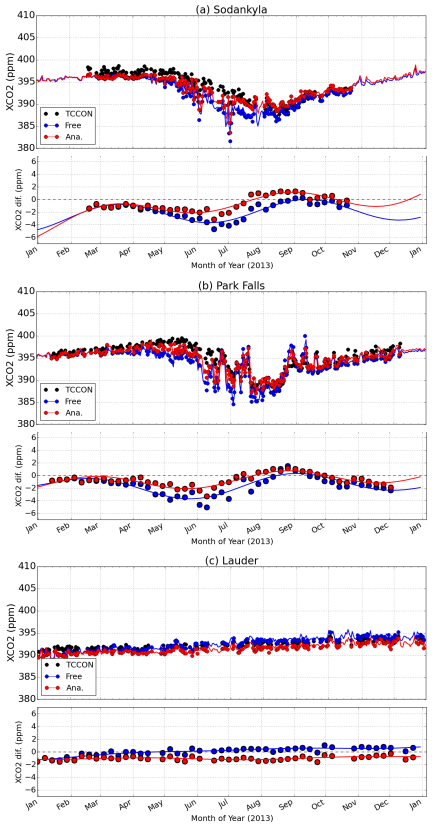

Figure 4. Time series of XCO_2 (in ppm) at **(a)** Sodankylä, **(b)** Park Falls and **(c)** Lauder. For each station, the top panel presents the daily averaged data from TCCON (black dots), the simulated XCO_2 (solid lines) and the daily averaged simulated XCO_2 in the observation space (coloured dots). The bottom panel presents the weekly averaged bias of the simulated XCO_2 against the TCCON data (coloured dots) and the smoothed bias (solid lines). The blue colour is for the free run while the red colour is for the analysis.

ACPD
15, 26273–26313, 2015

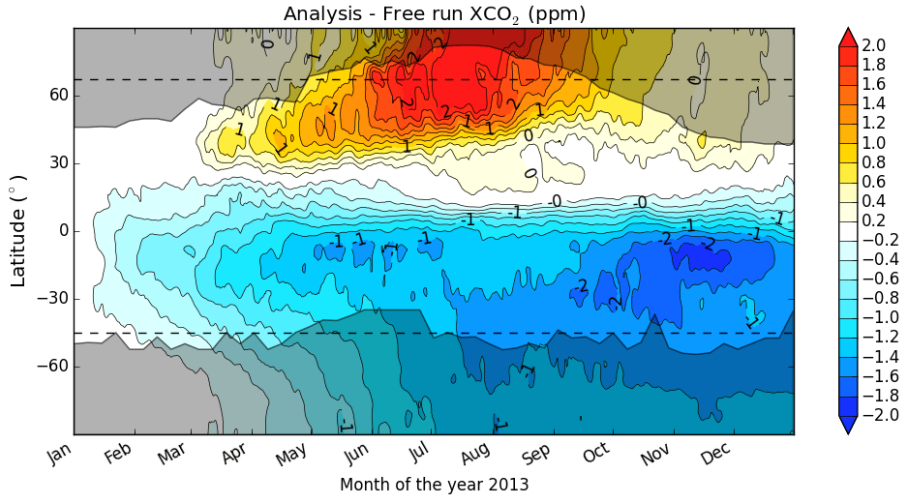

Figure 5. Hovmöller diagram (latitude vs. time) of the difference in ppm (negative/positive in blue/red) between XCO_2 from the analysis and from the free run simulation. The horizontal dotted lines represent the latitude of the northernmost and the southernmost TCCON station respectively. The grey shaded areas are where GOSAT does not provide observations.

ACPD
15, 26273–26313, 2015

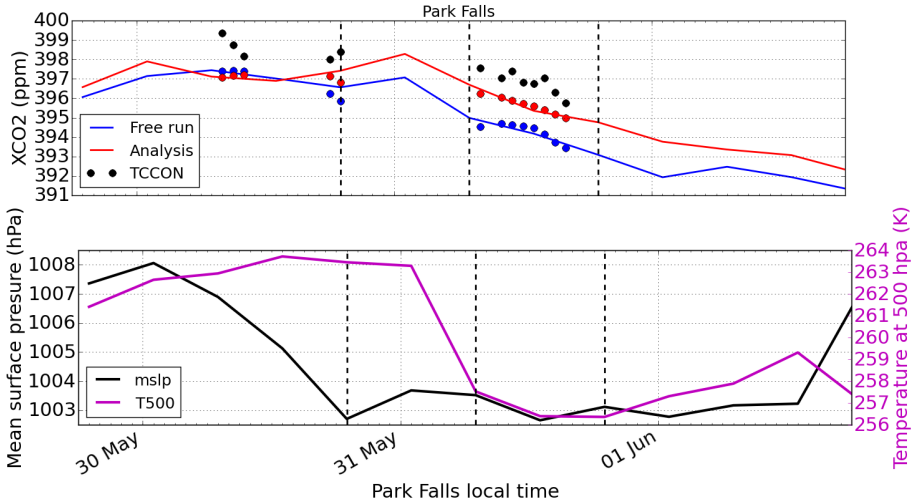

Figure 6. Situation over Park Falls between 30 May and 2 June. Top panel: evolution of XCO_2 (in ppm) from hourly averaged TCCON data (black dots), the free run (blue line and dots) and the analysis (red line and dots). The dots are the values of the model in the observation space. Lower panel: evolution of the mean sea level pressure (in hPa, black line) and the temperature at 500 hPa (in K, magenta line). The vertical dotted lines represent 31 May, at 00:00 UTC at 12:00 UTC, and the 1 June, at 00:00 UTC.

ACPD
15, 26273–26313, 2015

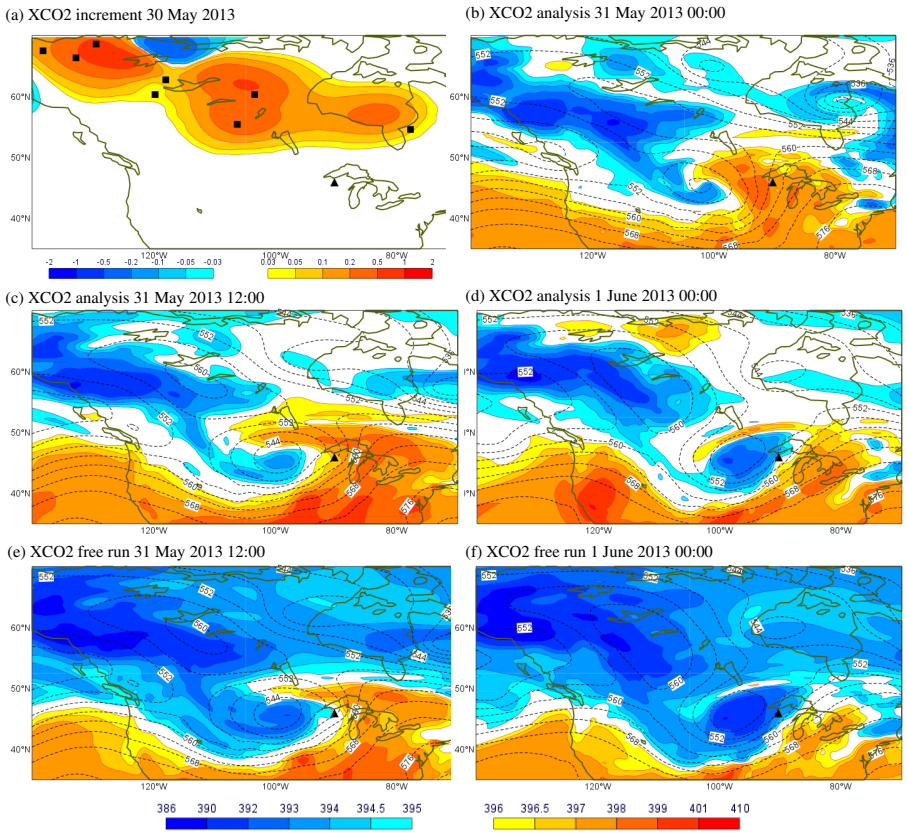

Figure 7. Situation around Park Falls (black triangle), Wisconsin, USA, end of May 2013. **(a)** average increment in terms of XCO_2 (in ppm) on 30 May 2013 (contours) and location of the GOSAT measurements during this day (black rectangles). **(b–d)** XCO_2 (in ppm) respectively on 31 May at 00:00 UTC, at 12:00 UTC and on 1 June at 00:00 UTC from the analysis. **(e, f)** XCO_2 (in ppm) on 31 May at 12:00 UTC and on 1 June at 00:00 UTC from the free run. For **(b)** to **(f)** the dark contours are the values of the geopotential at 500 hPa.

ACPD
15, 26273–26313, 2015

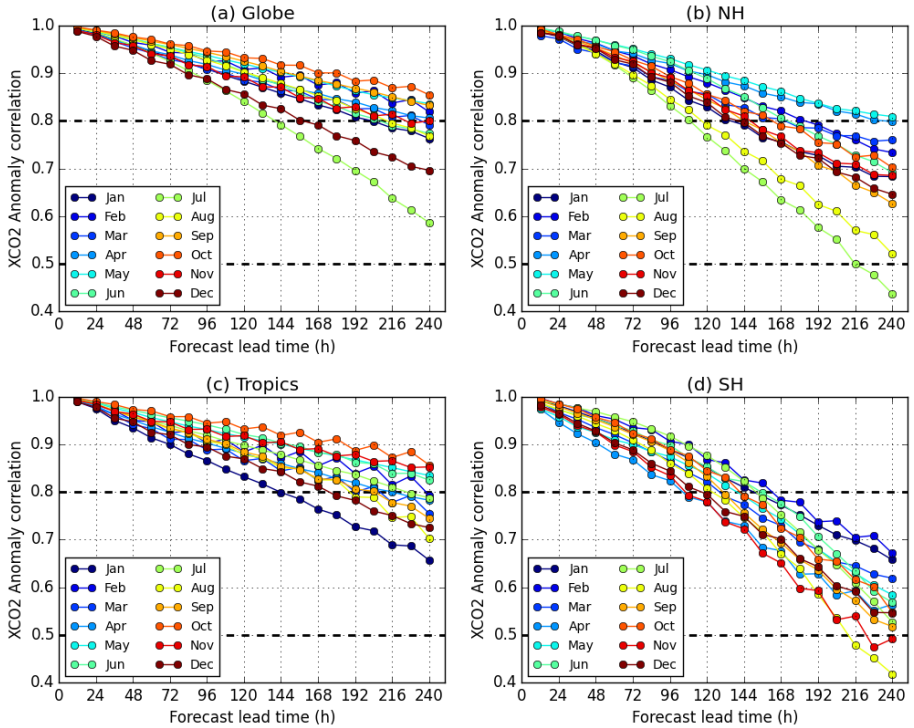

Figure 8. Anomaly correlation coefficient (ACC) of the forecast compared to its own analysis as a function of the forecast lead time and for each month: **(a)** global ACC, **(b)** ACC fot the Northern Hemisphere (20–90° N), **(c)** ACC fot the tropics (20° S–20° N), **(d)** ACC fot the Southern Hemisphere (90–20° S). Each month is represented by a different color (see inset legends).

ACPD
15, 26273–26313, 2015